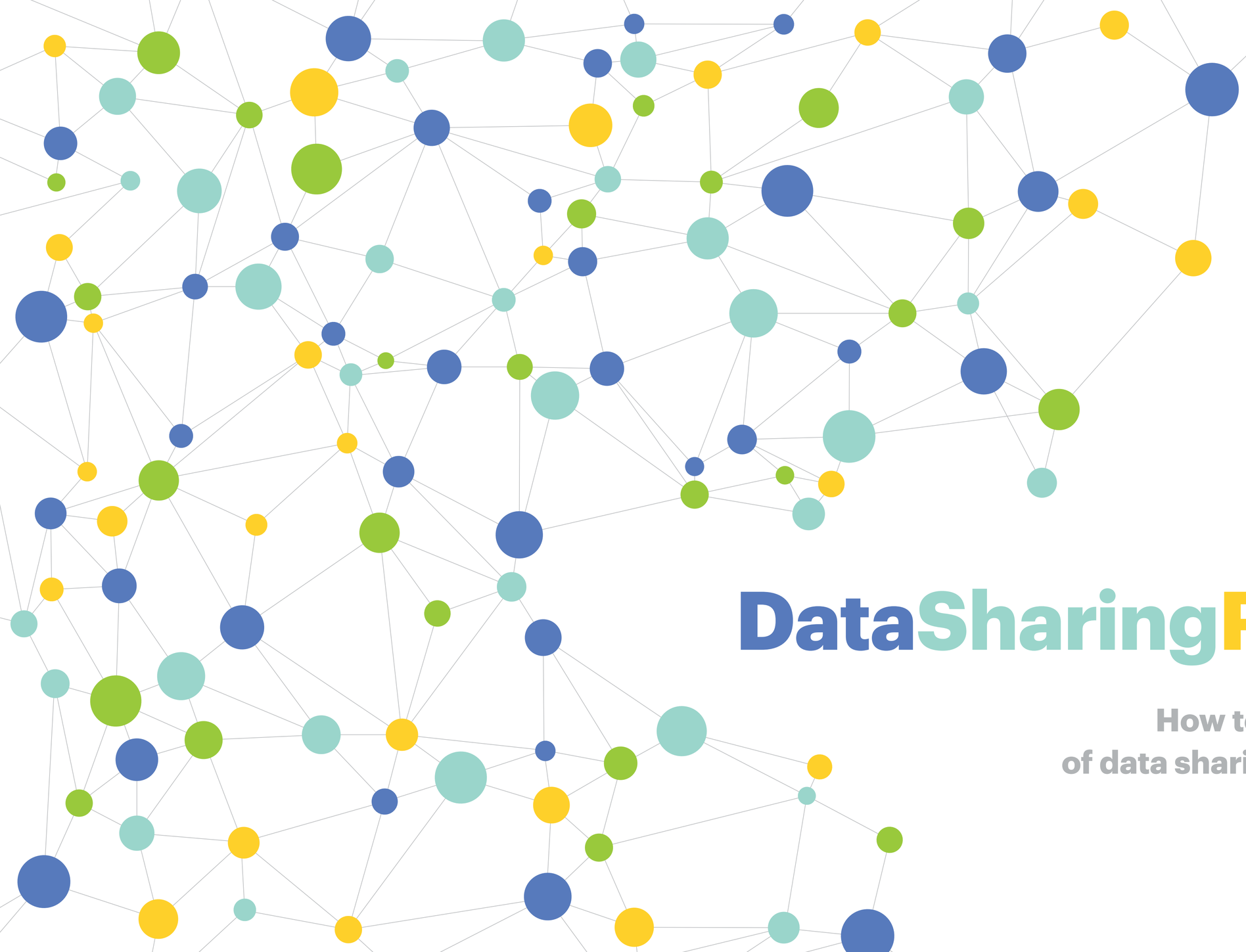


DataSharingPlayBook

How to unlock the potential of data sharing in IMI/IHI projects

The nature of many healthcare, science, and technology challenges and opportunities, increasingly calls for precompetitive multidisciplinary and public-private collaboration. The Innovative Medicines Initiative (IMI) and its successor, the Innovative Health Initiative (IHI), are pioneering this paradigm shift to advance health research and innovation. In this context, undoubtedly, data has taken a centre stage, playing a crucial role in boosting life-science research; it is also an area where the evolution of the precompetitive space and of culture —from IMI to IHI and more broadly the European Health Data Space— has been most visible.

Many of the IMI and IHI projects generate and/or make available to their partners and the research community an increasing volume and diversity of data. While doing so, all partners are regularly confronted with the same or similar questions related to legal, technical, data protection, ethical, and intellectual property challenges.

The purpose of this Playbook is to help all participants in pre-competitive projects by facilitating internal processes and decisions that accelerate the provision of data. This will create a positive impact on resources, project pace and overall collaboration spirit.

While addressing many of the industry processes, this Data Sharing Playbook offers solutions to all participants and gives the opportunity to better understand different perspectives when planning, executing, and exploiting results of data generated and shared in precompetitive consortia. By shedding light on data sharing workflows and processes, identifying common roadblocks and proposing standardised solutions based on best practices, successful past experiences, and available resources, this Playbook marks a turning point in unlocking the potential of data projects.

We encourage all involved parties to embrace the solutions suggested here, so that we can realise the power of innovation to change people's lives for the better.

Magda Chlebus

EXECUTIVE DIRECTOR SCIENCE POLICY AND REGULATORY AFFAIRS · EFPIA

Hugh Laverty

HEAD OF SCIENTIFIC OPERATIONS AT THE INNOVATIVE HEALTH INITIATIVE

TABLE OF CONTENTS

Tips to navigate the Playbook

From the **vertical menu** on the left-hand side, the user can move directly to any of the sections in the document. Once in a section, the corresponding colour circle in the menu expands to indicate where the user is in the Playbook. The **arrows on the right hand bottom corner** allow to swipe back and forth and return to a previously viewed page. By clicking on the **arrows at the top/bottom center,** the user can move to the previous/next page.

Within the swimlanes diagram, actions for each challenge are depicted on a horizontal axis that follows the various stages of the project. On this same page, it is also possible to filter by roles involved in each action.

Along the Playbook sections, the user will find **linked resources**. Resources, classified in five challenge areas, are easily accessible from relevant sections of the document. The complete list of resources is provided at the end of the Playbook.

R 2.3 – R 2.4

Main Menu

Go to next page

Go back/forth to previous page viewed

Why this Data Sharing Playbook?

Providing access to data for scientific research purposes is a common scenario in IMI/IHI projects. This has, in recent times, gained increasing importance and complexity. Organisations often contribute data to IMI/IHI consortia as an essential element of projects. However, the implementation of internal procedures, regulatory mandates, and technical processes to enable data provision represents a challenge. This often results in delays and inefficiencies which are tackled on a project-by-project basis. Standardised solutions for common obstacles could substantially boost the internal data sharing into IMI/IHI project partners and even beyond, increasing project efficiency and driving impact.

This Data Sharing Playbook provides strategies and resources to navigate common challenges associated with the provision of data in IMI/IHI projects, thus bringing efficiency into the process.

This Playbook presents data sharing challenges, together with strategies and solutions that facilitate decision-making to overcome those obstacles across the project life cycle. Consequently, internal resources are optimised, and experts can focus on what really matters: advancing collaborative and impactful research to improve health outcomes.

Who is it for?

Organisations with diverse roles and experience in data sharing projects were engaged during its development, to inform the list of requirements, identify common challenges, explore potential solutions and establish a common ground for best practices and project dataflows.

The Playbook was initially born out of a need from industry who had particular challenges. It has since been adapted to consider and be relevant for the broader community participating in IMI/IHI projects.

The Playbook is for all collaborators who are directly or indirectly involved in data sharing in IMI/IHI projects.

What to expect?

This Playbook intends to be a user-friendly and comprehensive tool for all those involved in data sharing in IMI/IHI projects. It outlines processes, challenges, and scenarios where roles, procedures, and resources are well depicted to inform, facilitate and accelerate data provision.

To this end, the Playbook provides insights into:

- Key data sharing **concepts** and main **roles/functions** involved.
- Key decisions to be taken during the project life cycle.
- **Root causes** of data sharing issues and **challenges** in different areas ***(PPP context, Legal & IP, Internal processes, Security, and GDPR).***
- **Strategies** to prevent and address these challenges which are then exemplified in best and worst-case scenarios.
- **Examples, templates, guidelines, publications, standards** and other useful **resources**.

This Playbook takes the user through the data sharing journey from the outline of the IMI/IHI topic and proposal preparation to the negotiation of the grant and consortium agreement and project execution. The post-project phase —sustainability and exploitation— is not the specific focus of this Playbook as it presents different challenges and scenarios, although several aspects of this Playbook will be applicable in this phase as well.

Executive summary

Data sharing is increasingly present in Health Research projects. While it enables more **impactful research for patients and citizens,** it also brings additional levels of complexity. In this Playbook we aim to untangle the complexity around data sharing in order to unlock and accelerate the unprecedented potential associated to the use of data in health research in the IMI/IHI context. The following have been identified as critical learnings:

- A **data sharing culture** shift is needed within organisations, where the added value of data sharing is understood and inner resistance is overcome. Old thoughts like *"this data is mine", "others do not know how to analyse or interpret my data"* should be substituted by acknowledgement of the power of data sharing to achieve better results, faster.

- Identify early on —ideally during the topic writing phase— which **type of data** will be needed to answer the project's research question/s (e.g., personal/non personal data).

- Create a **compelling business case** that showcases the added value that data sharing will bring for a specific project and for a participant organisation. Next, **assign value** to data as part of the partner's in-kind contribution. This will facilitate internal approval to share data.

- **Identify and involve the internal stakeholders.** Within an organisation, different roles and functions have a say to approve and expedite data sharing (i.e. academic leads, senior managers, statisticians, therapeutic area leads, lawyers, data protection officers). It is advisable to **engage them early** on to avoid obstacles downstream.

- Connect with experienced people within your organisation. The **IMI/IHI liaison officer** within each company may become a facilitator/aggregator of best practices and provide guidance to Principal Investigators (PIs).

- Within the consortium, decide on the **data sharing model** early on (centralised, federated, hybrid). The difference and consequences of this decision should be understood. Next to this, decide on the most appropriate **data anonymisation** strategy needed, for the data being provided for the project.

- The project **data flow/s** are a critical component to be depicted during the project preparation phase. It represents the transformations and flows that data will undergo in the project, the connections between those steps, and the stakeholders involved.

- Derive the **GDPR roles** based on the factual situation outlined in the data flow/s. Who will really determine data processing in the project? These partner(s) are the (joint) controllers.

- **Map the legal agreements** that will consequently be needed. Use pre-approved/previously used agreement templates whenever possible.

- **Check permission to use data**, be it ownership or license rights. It is essential to examine if informed consent is sufficient for the intended use. When collecting data prospectively, ensure that informed consent covers research use of data.

- Enter into a **dialogue with all consortium partners** to confirm there is a common understanding on the specifics of how data will be shared in the project. Involve consortium members in those discussions as soon as possible. This includes the development of the consortium Data Management Plan [DMP], a project deliverable due within the first 6 months of the project starting.

- **Frontload** the definition of data sharing aspects as much as possible. While not all matters can be defined before the start of the project, data sharing tends to be pushed to much later, thus creating bottlenecks, misunderstandings and causing potentially important delays.

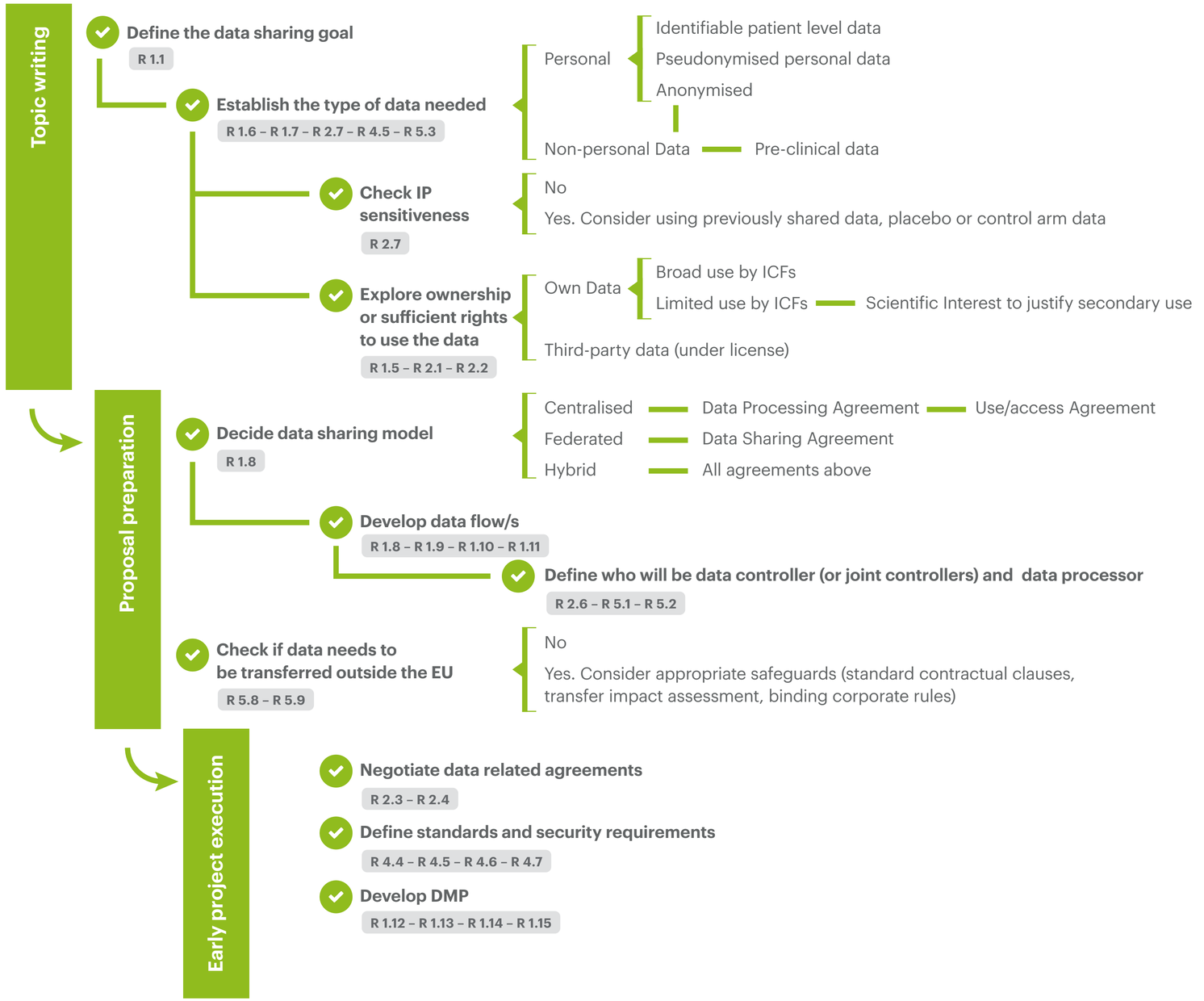
- For those aspects which cannot be fully decided, it is important to establish **early deliverables and milestones** to promote coming to conclusions related to data sharing aspects. The DMP, due by month 6 of the project, is a good document to include the pending data sharing details.

- The **security requirements** are often considered too late in the project. It is recommended to use accepted standards and agree to those early in the project to speed up reviews, as well as to use standard common data models and reuse of existing platforms and infrastructures rather than creating new ones.

KEY DECISIONS

Key decisions to unlock data sharing

This figure represents in a summarised manner key decisions to be taken during the project life cycle regarding data sharing. Links to resources available in the final section of the Playbook are offered next to each checkpoint.



Fundamental concepts

Identifiable patient level data
Information recorded on individuals that allows direct or indirect tracing to patients/participants. In a clinical trial context, this would include patient personal details, site, outcomes, etc.

Pseudonymised personal data R 5.6
Personal data that has gone through a 'pseudonymisation' process by which it can no longer be traced back to a specific individual (data subject) without the use of additional information. In most countries pseudonymised data is still considered personal data.

Pseudonymisation R 5.6
The processing of personal data in such a manner that it cannot longer be attributed to a specific individual (data subject) without the use of additional information. Such additional information must be kept separate and subject to technical and organisational measures. For instance, data allowing for direct identification are replaced with a code which is then stored in a separate location (i.e., a table). Data subject to Pseudonymisation is Pseudonymised Data.

Anonymised data R 5.3 – R 5.4 – R 5.5
Personal data that has gone through the process of removing sufficient information elements so that the individuals (data subjects) are not identifiable and cannot be re-identified by any means reasonably likely to be used (i.e., the risk of re-identification is sufficiently remote). Anonymous information is not personal data and data protection law does not apply.

Synthetic data
Data that has been artificially created by computer simulations and/or algorithms. These data mimic real-world data but without personal data or any IP-sensitive data elements.

Primary data R 5.7
Data directly collected from first-hand sources through different methods —surveys, interviews, experiments—, with a specific research purpose. Although collecting primary data is usually expensive and time-consuming, it ensures rights and control over the type of data to be generated and standards used.

Secondary use of data R 1.3 – R 5.7
Data previously collected and made available for others to use. This may include data generated by government, research, or healthcare institutions that are now used for a generic or different purpose than the one for which it was originally collected.

Federated hosting of data R 1.8
This model allows multiple distributed data sources to function as one. The federated network takes data from a range of sources that have been standardised to a common data model. This approach allows partners to query data from multiple sources. Yet, special attention needs to be paid to data standardisation, maintenance, and connectivity aspects.

Centralised hosting of data R 1.8
This model implies that the data are located, stored, and maintained in a single location. This approach often ensures the quality of the data, but costs, geographical location, and compliance with business and legal requirements can prove challenging.

Controller R 5.1 – R 5.2
A role defined in the GDPR as "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data". (i.e., the person or organisation that determines the purpose(s) for which personal data are processed and how such processing is to be done). A key role of the controller is to allocate responsibilities (i.e., who shall oversee compliance with data protection rules, and how data subjects can exercise their data ownership rights in practice).

Joint controller R 2.6 – R 5.1 – R 5.2
This role involves two or more organisations who jointly determine 'why' and 'how' personal data should be processed, which can be either by a common decision or by converging decisions. Since the responsibilities of joint controllers do not need to be equal, GDPR requires the joint controllers to clearly establish their respective responsibilities in what is usually called a joint controller arrangement.

Processor R 5.1 – R 5.2
A role defined in the GDPR as "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller". A processor may engage another processor (a "sub-processor") for specific tasks, but only after formal authorization by the controller. The mandate of the processor needs to be specified in a data processing agreement between the controller and the processor.

B

Background data

Any data that is held by an IMI/IHI project organisation before it accedes to the Grant Agreement, and which is needed to implement the project or exploit the results.

C

Clinical trial data (placebo & control vs. active arm data) R 1.2 – R 1.3

Clinical trial data refers, for the purposes of this Playbook, to data gathered during the conduct of a clinical study. Typically, clinical trials have at least two arms: the treatment arm and the control arm (alternative treatment/placebo arm).

Common Data Model (CDM) R 4.1 1 – R 4.12

A CDM is a way of organizing data into a standard structure. A CDM is useful for health care research because it allows for the harmonized pooling of patient information so that comparisons can be made about the relative effectiveness of different treatments. A CDM is particularly relevant when federated data hosting is used. OMOP is one such CDM.

Consortium Agreement (CA) R 1.1

The agreement signed between all partners (beneficiaries) in an IMI/IHI project. It supplements the provisions of the project Grant Agreement signed with the IMI/IHI. The CA specifies the rights and obligations of the project partners. For example, it contains provisions about the internal organisation and decision-making, financial matters, and the handling of intellectual property rights. It may also include mandates to the Coordinator and the Project leader, and pre-approved templates, for data and materials sharing.

D

Data Access Committee (DAC) R 1.16

A Data Access Committee is a body responsible for data release to external requestors based on consent and/or pre-agreed terms. Multiple datasets may be affiliated to a single DAC.

Data curation

The process of cleaning, harmonising, and organising datasets towards a common set of standards. It involves collecting, structuring, indexing, and cataloguing data.

Data Management Plan (DMP) R 1.12 – R 1.13 – R 1.14 – R 1.15

A DMP is a key element of good data management as part of making research data Findable, Accessible, Interoperable, and Re-usable (FAIR). A DMP describes the data management life cycle for the data to be collected, processed, and/or generated by a project. It is usually drafted as a deliverable in the first 6 months of an IMI/IHI project.

Data model R 1.8

Visual representation of the project's data elements and the connections between them. It also includes the definition and structure of the data needed to develop effective information systems. It enables organisations and technical resources to collaboratively decide how data will be stored, accessed, shared, updated, and leveraged across an organisation or project.

Data Transfer Agreement (DTA) / Data Sharing Agreement (DSA) / Data Access Agreement [DAA] R 2.3 – R 2.4

Contractual documents that set out the specific conditions under which data are made available for data sharing. They provide clarity to the receiving parties about their roles and responsibilities.

Data Protection Impact Assessment (DPIA) R 5.10

A DPIA is required under the GDPR whenever a new data processing activity is likely to generate a "high risk" to other people's personal information. It is a process designed to help systematically analyse, identify and minimise the data protection risks. The DPIA is usually drafted under the responsibility of the Controller.

Data Privacy Policy (DPP) R 1.4 – R 3.2 – R 3.3 – R 5.7

A DPP is a legal document that details the ways in which personal data may be used. At the very least, it needs to explain what data are collected, how data are collected, and how the collected data are processed.

Data Processing Agreement (DPA) R 5.1 – R 5.2

Legal agreement required by the GDPR between the controller and processor (i.e., a service provider). It regulates any personal data processing conducted by the processor and defines the mandate for the processor. If the processor wishes to involve sub-processors these should be defined in the DPA.

Data valuation R 1.1

It is the process by which industry partners assign a value to the data they provide to IMI/IHI projects in order to calculate their in-kind contribution.

Data workflow R 1.9 – R 1.10 – R 1.11

Sequence of tasks that processes a set of data. It depicts the phases or steps that data will undergo in the project, the connections between those steps and the stakeholders involved.

GLOSSARY

F

FAIR R 4.2 – R 4.3

This concept refers to the 4 principles that guide good data management: findability, accessibility, interoperability, and reusability.

G

General Data Protection Regulation (GDPR) R 5.1 – R 5.2

The General Data Protection Regulation (EU) 2016/679 is a legal framework that sets guidelines for the collection and processing of personal information from individuals who live in the European Union (EU). It establishes the regulations on data protection and privacy, and it also addresses the transfer of personal data outside the EU. The GDPR's primary aim is to enhance individuals' control and rights over their personal data and to simplify the regulatory environment for international business. It applies to any enterprise —regardless of its location and the data subjects' citizenship or residence— that is processing the personal information of individuals from inside the EU.

Grant Agreement (GA)

The GA is the funding agreement concluded between the IMI/IHI Joint Undertaking and the project's participants. It specifies the rights and obligations of the contracting parties and contains important provisions for the project implementation, such as criteria for the eligibility of costs and houses the full project plan called the Description of Action [DoA].

H

Honest broker / Trusted third party

An Honest broker is a neutral third party that collects confidential information from various parties and combines it to obtain aggregated information without disclosing the underlying confidential information. In most cases, honest broker systems are set up to obtain and provide clinical/medical records, data and specimens, but also for IP protection purposes.

I

Informed Consent Form (ICF) R 2.1 – R 2.2

An ICF is a document in which patients are given the information they need to make a decision, including possible risks and benefits, about a medical procedure, a treatment they will undergo, or a clinical trial they will participate in. It may include information on how data may be used for future research. The document needs to be signed independently by the patient. Consent forms are also commonly used to inform individuals (data subjects) and request their consent, about the processing of their personal data. These two types of consent forms are similar, and may even be combined, but are based on different legal frameworks (medical treatment acts/clinical trial regulation and GDPR respectively).

Intellectual Property (IP) R 2.7

Creations of the mind, project results such as inventions, literary, artistic works, designs, symbols, names, and images used in commerce. IP is protected in law by, for example, patents, copyright, and trademarks, which earn recognition or financial benefit.

J

Joint controller agreement R 2.6

A legal agreement to comply with the GDPR requirement (art 26) towards joint controllers to establish their respective responsibilities concerning personal data processing, in particular regarding the rights of data subjects. While joint controllers are jointly responsible, they still may have different and distinct roles that need to be transparently established in the joint controller agreement. Usually, the agreement also deals with mutual liabilities. Joint controller related provisions can also be made part of a data sharing agreement on consortium level.

L

Licensed data / third-party data R 1.3 – R 5.7

Data that is not owned by a project participant organisation, but which is accessible to it via a license with a third party. When sharing third-party data it is important to ensure that the user's rights attributed by the license are respected and communicated to the receiving party (e.g., through a Data Sharing Agreement).

M

Material Transfer Agreement (MTA) R 2.4

A contract between two organisations that governs the transfer of tangible research materials (such as biological samples) when the recipient intends to use them for its own research purposes. An MTA defines the conditions under which the materials are made available to the recipient. In many ways, it resembles the DTA, and in fact, often also encompasses a DTA when a rich data set is shared in conjunction with the materials.

R

Real-World Data (RWD)

Real-World Data (RWD) are data obtained from independent sources that follow healthcare outcomes in a representative population. RWD are observational and pertinent to actual clinical practice, in contrast to strictly controlled experimental data acquired in randomised clinical trials.

S

Standard Contractual Clauses (SCCs) R 5.8 – R 5.9

According to the GDPR, standard contractual clauses ensuring appropriate data protection safeguards can be used as a ground for data transfers from the EU to third countries. This includes model contract clauses — so-called Standard Contractual Clauses (SCCs)— that have been "pre-approved" by the European Commission. A transfer impact assessment, consisting of a written analysis of the impact that a transfer of personal data to a country outside of the EU may have on the privacy of the data subjects, needs to be conducted.

Roles

IMI/IHI liaison officer

The point of contact within an organisation (e.g., industry partners) for matters related to IMI/IHI projects. This role typically oversees cross-cutting aspects of the IMI/IHI project portfolio within an organisation.

Project leader

This person usually leads the project on the industry side. In charge, together with the project coordinator, of the overall scientific and project leadership.

Project coordinator

In IMI projects, this role leads the project on the public consortium side and is in charge, together with the project leader, of the overall scientific and project leadership. Additionally, this role oversees the grant administration aspects. In IHI this role can also be held by an industry partner.

Principal investigator

This role is responsible for the design and conduct of research (e.g., clinical trials), supervision of staff, producing deliverables, results and research findings. This role is the main contact person for scientific matters regarding a specific project.

Therapeutic lead

Responsible for leading and coordinating the organisation [usually industry] strategy regarding a disease area.

Senior manager/ Academic lead

Manager in a senior-level position who has the authority for planning and directing the work of a team, in a specific scientific area, and in line with strategic objectives and budget. This is a department leader role (or higher) and has responsibility to monitor the work of others and take corrective actions when necessary.

Data Protection Officer (DPO)

The guardian of data protection within an organisation. This individual acts as an independent advocate for the regulated care and use of personal information. This role is responsible for ensuring that an organisation is compliant with the GDPR and other relevant legislation. For some organisations, this role is mandatory by law. The role of a DPO is formally laid out by the EU as part of the GDPR.

GDPR expert

Specialist with a clear understanding of the General Data Protection Regulation (EU) 2016/679, a fundamental regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA).

IT specialist

Professional working in Information Technology within an organisation. The scope of this role comprises experts on computer programming, network administration, computer engineering, IT security, web development, technical support, and various other related tasks.

Statistician

Professional that specialises in the gathering, analysis, and interpretation of data to aid decision-making processes. This role participates in the design of studies and/or the statistical analysis of data.

Lawyer

This role provides legal advice, counsel and assistance within an organisation. This role may be part of a larger legal department with different expertise. More specifically, we can distinguish IP lawyers, privacy lawyers, regulatory attorneys, etc.

The description of these roles is indicative, reflecting the skills & expertise deemed necessary to facilitate data sharing.
All the roles are not necessarily involved in a project, depending on the objective of the project and the organization of each partner.

CH1

Public-Private-Partnerships (PPP) & Data Sharing Culture

INTRODUCTION & CONTEXT

It can be difficult to understand PPPs and the specific relationships between stakeholders, which are different and go well beyond the traditional customer-provider relationship. Specifically, this can be a challenge for newcomers to these types of partnerships; it can often be difficult to find agreement, appreciate the drivers for other parties, and both envisage and demonstrate added value to all involved. To demonstrate the value of PPPs it can be very helpful to highlight several parameters from impactful PPPs, for example, regulatory impact, improved endpoints, prediction biomarkers, patient segmentation, technology tracking, and insights gained from analyses of big data.

FREQUENTLY ASKED QUESTIONS

CHALLENGES

How to best construct IMI/IHI consortia to ensure internal support that enables effective and efficient data sharing?

- Consider strongly the involvement of company experts with PPP experience in proposal drafting.
- Conduct PPP training for novice project leads and coordinators to gain important insights at both the operational and senior levels.
- Create a network(s) of IMI/IHI project leads and coordinators inside and across organisation partners that can share experiences and best practices.
- Keep in mind the sustainability of project outputs.

RESOURCES

1.3 Sharing and reuse of individual participant data from clinical trials: principles and recommendations
1.4 Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis
1.5 Sharing Is Caring – Data Sharing Initiatives in Healthcare
1.6 Data sharing policy: example of EOSC-Life Data Sharing Policy of the COVID-19 repository
1.7 BigData@Heart - Responsible data sharing in a big data-driven translational research platform: lessons learned
3.5 Involvement of Roles in Data Sharing Decisions

How to best extract, define and ensure added value from PPP?

- Construct a clear and easy-to-apply methodology to assign value to data sharing in order to stimulate the sharing in IMI/IHI projects.
- Report and disseminate impact(s) of previous PPP projects as an exemplar for the present project.

RESOURCES

1.1 The case of data sharing in precompetitive settings
1.2 Status, use and impact of sharing individual participant data from clinical trials: a scoping review

How to best deduce and understand PPP data flow(s) within IMI/IHI projects?

- Create and assess the entire project data flow, including 'who has access to what?' before the project starts.
- Identify the key players so that they assist in the data flow preparation, supported by a consequent workplan.
- If this cannot be done at such an early stage, plan to agree and establish the data flow/s, as part of the DMP.

RESOURCES

1.6 Data sharing policy: example of EOSC-Life Data Sharing Policy of the COVID-19 repository
1.8 Schematic data flows with GDPR roles
1.9 Dataflow overview
1.10 IMI HARMONY Data Flows
1.11 IMI H2O Data Flows

How to best avoid commonly seen pitfalls related to data sharing in PPP?

- Confirm with all involved partners (i.e., in a workshop setting) that they understand and agree with the committed level of data sharing (not only high-level principles but what it means exactly for each partner) as early as possible.
- Ensure that foreseen capacities are clearly described and budgeted.
- Consider procurement of services that are not the core business of any of the partners in the consortium, as these tend to become bottlenecks otherwise.

RESOURCES

1.12 Data Management in EU Collaborative projects (not yet available for HE)
1.13 DMP template Horizon Europe
1.14 DMP IMI H2O Project
1.15 DMP IMI Conception Project
1.16 Data Access Committee

CH2 Legal/IP
CH3 Internal Processes
CH4 Security/IT
CH5 GDPR

CH2

Legal & Intellectual Property (IP)

INTRODUCTION & CONTEXT

Agreeing on the legal framework and contracts within PPP can be incredibly time-consuming. Although contracts and frameworks are often similar, there is no clear vision on how all legal components/agreements link together. Therefore, establishing agreements across large consortia and accompanied procedural scrutiny are often lengthy processes. Involvement of legal experts with PPP experience (i.e., previous projects) in proposal drafting is often beneficial but due to workload of these seasoned personnel, this is commonly not easy to achieve.

FREQUENTLY ASKED QUESTIONS

How can partners best select which data to share in PPP projects?

- Consider protection of IP in PPP, in particular when compound-related data are involved.
- Check ownership and third-party rights on the data before making any commitments for data sharing.
- Consider sharing data that have already been shared and approved (i.e., in previous PPP and/or with regulatory authorities). A new approval process will likely be needed (due to a different purpose of use for the data), but the fact that it has been shared before may facilitate the ability to share it.
- Assess as early as possible the confirmed ability, in terms of consent and ethical approval, to share data that are planned to be shared.

RESOURCES

1.4 Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis

1.5 Sharing Is Caring – Data Sharing Initiatives in Healthcare

2.7 How does Pharmaceutical IP work?

How best to check consent agreements for data within large PPP consortia?

- When checking for “informed consent” bear in mind that consent for participation in a trial and consent for access to data (GDPR) are in principle two different things. Clinicians will often, in first instance, think about the former definition.
- Confirm as early as possible that consent agreements and/or licenses allow the sharing of data that are planned to be used in the PPP in order to avoid potentially lengthy delays. Checking of the ICF language can be done by a regulatory attorney, privacy officer or CT manager. If there is a need to adapt the study ICF template with project specific wording (in case of prospective studies performed as part of the project or as in-kind contributions), then the regulatory attorney will always have to be involved.
- Request that all partners in a PPP make available their Informed Consent Forms (ICFs) templates to check the permissibility of intended data usage within the PPP. If ICFs templates cannot be shared, specify what language needs to be included in the ICF to enable use of data for the project. Then each partner can incorporate such additional language in their own templates. In these situations, where the ICF cannot be shared, it is recommendable to contractually capture by way of representations and warranties of the contributor that compliant ICFs are in place, that they allow for implementing project tasks and that state which restrictions are applicable for Research Use. The use restrictions from the ICF can also be formulated in a data intake form for the project, e.g., a Terms of Use (ToU) form.
- Be aware that already existing data sets often have incomplete informed consent in terms of secondary use which may complicate further data sharing considerably.

RESOURCES

2.1 IMI DO-IT Project Informed Consent Forms templated

2.2 German Medicine Informatics Initiative broad consent

How to construct effective legal frameworks in PPP consortia efficiently?

- Ensure every consortium partner legal colleagues are actively involved as soon as possible to facilitate understanding and the sense of collaboration in PPP.
- Construct a clear idea of the data flow/s in the project before the Consortium Agreement (CA) is signed. The data flow will be essential in determining the legal agreements needed in the project. Start with a high-level understanding and define clear milestones/deliverables early in the project to determine the full details.

CHALLENGES

CH1 PPP and Data Sharing Processes

CH3 Internal Processes

CH4 Security/IT

CH5 GDPR

CH2

Legal & Intellectual Property (IP)

- Grow the legal framework over the project lifetime: start with everything needed for project use, then internal research use, and then external use.
- Establish common definitions in linked/connected agreements in order to avoid inconsistencies.
- Use standardised data sharing templates to avoid lengthy detailed discussions.
- Develop a map of how components in a PPP legal framework link together.
- Consider mandates that permit (e.g., data sharing agreements for the PPP) to be signed by only the project coordinator and/or project leader, as long as the parties have agreed first to a template agreement and no relevant changes are made when using such template. An example could be the scenario of Joint Controllership between Beneficiaries for all the contributed data. It should be avoided that reoccurring legal agreements with third parties are mandated to be signed by all project partners.
- Consider the possibility of initially sharing data only with a subset of the consortium to facilitate rapid agreement and an early start (e.g., at the Work Package level). Having all partners involved in agreeing on each contract can be very time-consuming.
- Consider bilateral MTA/DTA as an alternative, although this is not an optimum strategy to overcome the burden of having many partners negotiating an agreement. This should not be used to exclude data or results from the project but can be an effective strategy to kick-start certain project activities as a temporary fix whilst the wider agreement is being negotiated. This allows partners to start working on the data/materials under a bilateral contract to prevent delays in performing the Action. Once the consortium DSA is in place, this will replace the bilateral ones.
- Include additional annex(es) to the CA to increase efficiency in data sharing within PPP. For example, it is possible to include a data sharing agreement template as an appendix to the CA either at the time of signature or thereafter via an Amendment to the CA.

RESOURCES

2.3 IMI2 Data Transfer Agreement, including User Confirmation Sheet
2.4 IMI EPND Material and Data Transfer Agreement
2.5 Anonymised example of Data Access Request Form
2.6 Joint Controllership Agreement
3.5 Involvement of Roles in Data Sharing Decisions

CHALLENGES

CH1 PPP and Data Sharing Processes
CH3 Internal Processes
CH4 Security/IT
CH5 GDPR

CH3

Internal Processes

INTRODUCTION & CONTEXT

Unclear internal data release processes and chains of decision-making can often prevent/delay the sharing of data in PPP, leading to long lead times between the decision to share data and actual data sharing.

FREQUENTLY ASKED QUESTIONS:

How to best expediate data sharing within PPP?

- Identify and involve all internal stakeholders as soon as possible in the discussions and ensure their buy-in.
- Assess data sharing in a risk/reward analysis with clear details on the intended data usage(s).
- Consider using demonstrators of impact(s) from other PPP data sharing as helpful and persuasive reference.

RESOURCES

1.1 The case of data sharing in precompetitive settings
3.4 Frontloading data sharing decisions in IMI/IHI projects: advantages and limitations
3.5 Involvement of Roles in Data Sharing Decisions

How to best establish internal data release processes?

- Establish organisational policy on data sharing.
- Prepare as early as possible key decision makers for the required sign-off to approve data sharing in PPP.
- Identify and engage early in the process internal stakeholders that need to be involved in data sharing approval. Several functions may need to be consulted. Ask them to review all critical documents, such as the ICF and the DMP.
- Highlight the criteria for data release decision in the internal guidance. It should be clear who has the authority to sign-off on data release.
- Include a specific internal IMI/IHI liaison role. This role is instrumental in providing guidance on how to navigate internal data sharing clearance. Involve the legal departments in discussions and project planning.
- Include internal data owners in the planning of data sharing in order to incorporate their perspectives and avoid eventual issues/blockades.
- Consider creating/adapting internal tools to expedite internal review and approvals of data releases. This can result in significant time savings.

RESOURCES

3.1 Testimonial on internal company tool to facilitate release of data by Sean Turner
3.2 Data Sharing Policies of Vivli members
3.3 MSD Portal and procedure to Access to Clinical Trial Data / Public website and document
3.5 Involvement of Roles in Data Sharing Decisions

How to best select and identify which data to share in a PPP?

- Be sure to secure sufficient time in the project planning stages to discuss data sharing.
- Take into consideration the decision-making timelines within companies, as these will be time-consuming.
- Determine required data quality and selection processes as early as possible.
- Plan early triage of data. This can assess data quality, completeness, and/or standardisation level(s) to ensure that data match pre-defined criteria.
- Plan the resource efforts needed to ensure that data criteria are adhered to by consortium members.
- Consider and keep in mind the sustainability of data shared in a PPP, and the additional data generated in any project.

RESOURCES

1.5 Sharing Is Caring – Data Sharing Initiatives in Healthcare

CHALLENGES

CH1 PPP and Data Sharing Processes | CH2 Legal/IP | CH4 Security/IT | CH5 GDPR

CH4

Security & Technology

INTRODUCTION & CONTEXT

Ensuring that required IT security criteria are met in PPPs can lead to delays in data sharing. The GDPR and increased digitalization mean that this may be perceived as less of an issue nowadays, but the focus on other areas also implies that the security requirements are often considered only very late in the project. Due to a lack of agreed IT and/or security standards and differing perspectives in consortia, IT security reviews often identify long lists of risks to be addressed.

FREQUENTLY ASKED QUESTIONS:

What [technical] environment will this PPP use to share data?

- Establish as early as possible (i.e., already in the grant proposal stage) the PPP data sharing concept at a high-level (i.e., central vs federated data storage).
- Consider open-source solutions that might be beneficial compared to proprietary solutions because of potential lock-in issues and/or lack of interoperability.
- Consider utilising a well-established and standardised Common Data Model(s) and platform(s) to enable data sharing and subsequent data usage(s). Budget for the data transformations needed to arrive at that model.
- Consider the reuse of proven solutions from previous PPP, rather than creating completely new platforms. The overhead involved in establishing legal frameworks, security reviews, etc. should not be underestimated.
- Consider early what analyses will be performed on the shared data within a PPP as this will have a strong impact on environment architecture decisions.

RESOURCES

3.5 Involvement of Roles in Data Sharing Decisions
4.1 IMI Criteria for sharing data
4.2 FAIR data principles
4.3 FAIR cookbook

What are the IT security requirements for this PPP?

- Consider IT security requirements based on well-established common standards (i.e., ISO) in parallel to legal and IP requirements. This could avoid lengthy and detailed reviews.
- Consider adding IT security requirements for transfer, hosting, and access as an annex to the Consortium Agreement (CA) or in the Data Management Plan (which is already an annex to the CA).
- Align IT security requirements in PPP and nominate one partner to conduct the IT security review for all in order to reduce overheads and/or delays.
- Discuss and agree the PPP data anonymisation strategy as early as possible (i.e., anonymisation, pseudonymisation, other). Also, be aware of significant differences in pseudonymisation standards between countries.
- Consider applying approved IT security standards. If this is not the case, the reasons for deviation should be clearly described and explained.

RESOURCES

4.1 IMI Criteria for sharing data
4.2 FAIR data principles
4.3 FAIR cookbook
4.4 ISO Information Security standards
4.5 ISO/TC 215 Health informatics
4.6 CDISC formatting for structured data
4.7 EOSC-Life Report on data standards
4.8 SEND format for harmonized structured preclinical data
4.9 Fast Healthcare Interoperability Resources (FHIR)
4.10 Health Level 7
4.11 OMOP Common data model
4.12 EHDEN 101: What is a federated data network? What is the OMOP common data model?
4.13 External guidance on the implementation of the European Medicines Agency policy on the publication of clinical data for medicinal products for human use
4.14 Health IT Standards

CHALLENGES

CH1 PPP and Data Sharing Processes
CH2 Legal/IP
CH3 Internal Processes
CH5 GDPR

CH5

General Data Protection Regulation (GDPR)

INTRODUCTION & CONTEXT

Concerns related to the GDPR within PPP's are frequent and can lead to significant issues concerning data sharing. Implementation of the GDPR has many country-specific, and company-specific interpretations which results in several different perspectives within almost all PPPs. International transfers outside the EU pose specific issues, relevant for companies with sites outside the EU (mostly in the US). In general, it is considered best practice to check upfront the alignment of Informed Consent Forms (ICF) with GDPR; the ability to share data inside and outside of the consortium; sharing of data geographically; the permission for secondary use of data in a project; and the applicability of data to the intended research purposes.

FREQUENTLY ASKED QUESTIONS:

What are the key best practices for addressing how the GDPR applies to a PPP?

- Consider that all PPPs with an important data sharing component must have an ethical/legal framework, an expert legal partner on board, and all partners sharing personal data should nominate a responsible person for data protection matters.
- Ensure that data protection officer(s) of the organisations are involved in the data sharing discussions.
- Consider using the same concepts and definitions in PPP agreements and documents as are described directly in the GDPR. Potentially these can be included as part of the PPP Data Management Plan.
- Map roles and accountabilities in the PPP data flow early on, in order to establish the data sharing concept.
- Establish roles and responsibilities in binding legal contracts (controller to processor agreements or joint controller agreements). The GDPR roles should reflect reality rather than the self-perceived roles in the project. This requires a detailed data flow analysis. The outcome could be that multiple parties are actually jointly responsible for the personal data (Joint Data Controllers). These agreements may be part of the Consortium Agreement.
- Consider that GDPR only applies to personal data. It is possible to share only fully anonymised data to ease data sharing. However, this can be limiting to the application(s) of data within projects.

RESOURCES

1.8 Schematic data flows with GDPR roles
3.5 Involvement of Roles in Data Sharing Decisions
5.1 European Data Protection Board Guidelines on the concepts of controller and processor
5.2 Key roles in GDPR
5.3 Introduction to anonymisation
5.4 IMI HARMONY Anonymisation Concept
5.5 Sharing Anonymised and Functionally Effective (SAFE) Data Standard for Safely Sharing Rich Clinical Trial Data
5.6 Report on Deploying Pseudonymisation Techniques by ENISA (European Union Agency for Cybersecurity)
5.7 TRANSCELERATE - A Privacy Framework for Clinical Data Reuse: Secondary Data Use in the Pharmaceutical Industry
5.10 Data Protection Impact Assessment

How to share data outside of the EU in compliance with the GDPR?

- Establish the necessary terms in standardised common and broad ICF template utilized in the PPP (only possible if data are actively collected; not for secondary use).
- Consider that, when sharing with non-EU sites of partners, pragmatically it can work best when the European site is the party receiving data access under the CA; the internal corporate binding contractual clauses can then be used for sharing onwards with the foreign sites.
- Implement a 1:1 contract between the data host or the data contributor and non-EU partner rather than consortium-wide agreements.
- Consider that recent legal rulings (SCHREMS II) severely limit the possibilities to share personal data outside the EU, but standard contractual clauses are still possible under SCHREMS II ruling. However, this requires an assessment of the local data protection of the foreign sharing partner (transfer impact assessment). SCCs can form part of the Data Sharing Agreement.

RESOURCES

5.8 The CJEU judgment in the Schrems II case by European Parliament
5.9 SCHREMS II Summary

CHALLENGES

CH1 PPP and Data Sharing Processes
CH2 Legal/IP
CH3 Internal Processes
CH4 Security/IT

SCENARIOS

Scenarios

This section presents a fictitious case of an IHI project involving data sharing. The left column represents recommended choices and considerations in handling data sharing. In the right column, a scenario is given where best practice is systematically not followed in order to illustrate the consequences. Next to these, the links to relevant strategies presented in the Challenges section are provided.

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

The call text requires to include an ethical/legal work package.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

- ✔ Seasoned scientists (with a track record of successful IMI projects) from the participating companies involve senior management and also data stewards of their organisation(s) to define clinical trial data that could be committed for data sharing.
- ✔ A long list of potential company datasets to be shared is defined in this manner.
- ✔ A stakeholder analysis is performed within each company to determine who needs to be consulted internally for data sharing.
- ✔ The first legal checks (existing data licenses, IP) on the availability of the proposed datasets are performed.
- ✔ The call text highlights that the clinical data from companies should be integrated with academic real-world data in a joint central database securely managed by a qualified partner. Additional other data sources will be connected in a federated set up, based on a common data model.
- ✔ A clear business case is developed for each of the industry partners to establish the benefits of participation.

Worst case scenario

- 🚩 Enthusiastic scientists with no prior IMI project experience (experts in the disease topic) from the participating companies commit to the contribution of clinical trial data without consulting the trial leads. The call text highlights that the clinical data from companies should be integrated with academic real-world data in a joint central database securely managed by a qualified project partner.

Challenges FAQS and strategies

CHALLENGE 1

How to best construct IMI/IHI consortia to include data experts and ensure organisational support in order to enable effective and efficient data sharing?

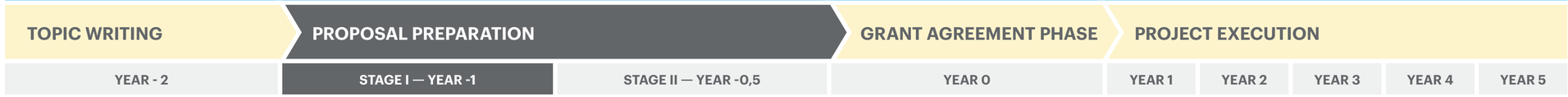
CHALLENGE 2

How to construct effective legal frameworks in PPP consortia efficiently?

CHALLENGE 2

How best to check consent agreements for data within large PPP consortia?

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.



Best case scenario

- ✔ The academic partners submit a scientifically compelling proposal with a broad range of clinical sites on board that could bring in observational clinical data from clinical practice. The commitment is supported by all the organisations, clinical leads as well as data stewards and DPOs.
- ✔ A specific Ethical, Legal and Societal Issues (ELSI) work package has been written, led by an experienced (legal) research group who has successfully dealt with this aspect in a previous large European project with partners in different countries.
- ✔ A data hosting partner has been included with previous experience in hosting and processing personal data with track record in central as well as federated settings. Previous IT audits were successfully conducted.

Worst case scenario

- The partners submit a scientifically compelling proposal with a broad range of clinical sites on board that could bring in observational clinical data from clinical practice without checking Informed Consents or having relevant ethical approval. Moreover, there is no ethics clearance that previously collected clinical data can be reused.
- A specific ethical/legal work package has been written, led by one of the partners who has successfully dealt with this aspect in a previous European project with 4 partners.
- A data hosting partner has been included with previous experience in hosting personal data in its own country but has never been subject to an audit.

Challenges
FAQS and strategies

CHALLENGE 1

How to best construct IMI/IHI consortia to include business matter experts and ensure executive power in order to enable effective and efficient data sharing?

CHALLENGE 2

How to construct effective legal frameworks in PPP consortia efficiently?

CHALLENGE 2

How best to check consent agreements for data within large PPP consortia?

SCENARIOS

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

- ✔ The partners come together in the consortium; scientifically there is a great match and enthusiasm about the wealth of data available. There are two data leaders to cater against staff moving out during the project.
- ✔ The selection criteria for the data sets to be shared within the project are agreed with the consortium.
- ✔ At a high-level, the data sharing concept between the organisations and the data hosting partner appears to be matching. The flow of data including the use and access criteria is being described in some detail. A comprehensive list of data sources is available indicating data formats, volume, ethical approval confirmation and key contacts. Drafts for Data Processing and Data Sharing agreements are circulated with legal colleagues.
- ✔ The IT departments of companies are involved and have nominated one party as the representative defining the security requirements for data sharing; this is defined as a deliverable for project month 12.

Worst case scenario

- The partners join the consortium and scientifically there is a great match and enthusiasm about the wealth of data available.
- ⚑ At a high-level the data sharing concept between the organisations and the hosting partner appears to be matching, but on the details, there are incompatible data sharing concepts proposed. It is decided to deal with those in the first 6 months of the project in the Data Management Plan.

Challenges
FAQS and strategies

CHALLENGE 1

How to best avoid commonly see pitfalls related to data sharing in PPP?

CHALLENGE 2

How to construct effective legal frameworks in PPP consortia efficiently?

SCENARIOS

SCENARIOS

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

- ✔ During the grant agreement phase the lawyers (who did the initial legal check) join the discussion and necessary definitions requested by the GDPR are finalized (i.e., controller/joint-controller/processor roles) in light of the available data flows for central or federated set-ups. There is agreement about the GDPR roles. The hosting partner is willing to take up the controller role, but there is a fair settlement of the liabilities between the consortium partners in the Joint Controller Agreement.
- ✔ Use and access are being clarified as well as background and foreground IP questions. Potential data users outside the EEA are identified and data is accessible also for them.
- ✔ The legal data sharing agreement first draft has been agreed and aligned early with the key data leads of the associated WPs and the first draft circulated to consortium legal contacts.
- ✔ Any questions regarding potential third-party rights on the data to be shared are being solved.
- ✔ An outline of the DMP is developed, including key components of data management and legal and GDPR aspects.

Worst case scenario

- 🚩 During the grant agreement phase the partner's lawyers join the discussion for the first time and demand clarity on the GDPR controller/processor. It is proposed to copy a legal set-up from a previous IHI project where the data hosting party took full controller responsibility. Other parties do not want to commit to transfer of controllership and the data hosting party is not willing to take up the liabilities of becoming a GDPR controller. The leader of the legal/ethical WP is not able to address all these questions and to lead the discussions due to lack of in-depth legal expertise.
- 🚩 It turns out to be impossible to devise an alternative division of responsibilities in the few months left during the grant agreement phase due to lack of clarity on the exact data sharing concept (still to be worked out in the DMP).
- 🚩 It was therefore decided to delay this decision untill the project execution phase.

Challenges FAQS and strategies

CHALLENGE 5

What are the key best practices for addressing how the GDPR applies to this PPP?

CHALLENGE 2

How to construct effective legal frameworks in PPP consortia efficiently?

SCENARIOS

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

- ✔ The first version of the DMP is finalized in the first 6 months of the project. It provides the detailed dataflows, ID management, infrastructure descriptions (secure data environment), details on use and access and confirms the previously agreed on GDPR roles.
- ✔ Project management is requesting the first test datasets from the partners involved. Data transfer (central) and data access (federated) works as described and confirms that the data flow diagrams are complete and correct.
- ✔ The earlier established business case was used to convince internal stakeholders of the benefits for data sharing. Data transfer is done using secure transfer modes (not via mail).
- ✔ Organisations follow their standard internal clearance process for data release.
- ✔ Some datasets turn out to be unavailable for data sharing due to lack of informed consents, but this does not cause issues since there are many data sources being contributed.
- ✔ Some datasets with only control arm data were selected as first datasets to be used as pilot data. The release of these datasets turned out to be straightforward.
- ✔ The security requirements are agreed between the data hosting partner and the company taking the lead in defining these. This is described in a project deliverable.
- ✔ The data transfer method is agreed as part of the IT security discussions.

Worst case scenario

- 🚩 A DMP is drafted in the first 6 months of the project using the standard template available. It provides high-level dataflows, but it was not possible to work out all the details due to lack of time and available manpower (hiring staff proves to be slow).
- The data hosting partner provides the first pilot installation of the secure data environment.
- The ethical/legal partner is proposing that the data contributing partners are data controller and the data hosting partner will the data processor.
- 🚩 The contact persons of the data contributing partners are trying to get help from their legal departments. Some are unable to locate anyone to help, others need to file an ethical approval but don't have the appropriate documentation from the project, others get very specific requirement from their lawyers regarding the pseudonymisation of the data. It turns out that the requirements are very different depending on the local regulations despite having one European GDPR.
- Project management is requesting the first test datasets from the partners involved.
- 🚩 Legal departments are involved and like to know: do partners know whether there are any third-party rights on the data? What agreements are in place with the data hosting party? Will the companies IP still be protected when sharing real clinical trial data? It takes the consortium on average 6 months to deal with these first round of questions. Some of the datasets originally selected were actually in-licensed from other sources and could not be shared at all.

Challenges FAQS and strategies

CHALLENGE 2
How to construct effective legal frameworks in PPP consortia efficiently?

CHALLENGE 4
What are the IT security requirements for this PPP?

CHALLENGE 5
What are the key best practices for addressing how the GDPR applies to this PPP?

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

✔ The conditions for further use of the data by the beneficiaries outside the scope of the project ("internal research use") are established. This requires an update of the consortium agreement, which is signed by the end of the year.

Worst case scenario

- Negotiation of the data processing agreement as part of the Consortium Agreement turned out to be cumbersome but was finally signed by all partners.
- 🚩 Getting approval for the first partner datasets proved to be very challenging, while the details of the project were not clear for most of the others after so many years.
- 🚩 The new partners of the IHI project were not able to provide the business case rapidly as they were not involved in the initial discussions around the call definition 4 years ago. This hurdle was overcome in most companies after about 6 months.
- 🚩 When trying to get access to the data, the responsible partners block the release of the data since they weren't consulted when selecting the data sets nearly 5 years ago. They feel responsible for the data and are not willing to share the data externnally.

Challenges
FAQS and strategies

CHALLENGE 1
How to best construct IMI/IHI consortia to include business matter experts and ensure executive power in order to enable effective and efficient data sharing?

CHALLENGE 1
How to best extract, define and ensure added value from PPP?

CHALLENGE 1
How to best select and identify which data to share in a PPP?

CHALLENGE 3
How to best expediate data sharing within PPP?

CHALLENGE 3
How to best establish internal data release processes?

SCENARIOS

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

- ✔ One of the project leaders left the organisation in order to pursue a new challenge. She was replaced by an enthusiastic young colleague who could be brought up to speed easily by the other leader.
- ✔ One of the IT departments decided to perform a security audit at the data hosting partner. Thanks to the clear description of the security requirements the hosting party could easily demonstrate compliance.
- ✔ The first dataset was shared with a partner outside the EU using an agreement between the data host and this non-EU partner.

Worst case scenario

- After lengthy discussions it was agreed to at least release the control arm of the data.
- 🚩 The IT staff of the first organisation contacted the data hosting party and requested an audit of the security measures in place. It turns out the information was only available in the local language, and a national IT auditing standard was used which was not recognized by the company IT staff.
- After 6 months of discussions, the organisation and the data hosting party agreed on a number of additional control measures and the English-language documentation required.
- 🚩 The organsation prepared the first dataset for transfer and requested the preferred methods of data transfer. The data hosting organisation did not have a standard procedure for that. After a few weeks of e-mails going back and forth a data transfer method was agreed which was implemented 3 months later.

Challenges
FAQS and strategies

CHALLENGE 4

What environment will this PPP use to share data?

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

- ✔ The updated Consortium Agreement was signed by all partners allowing data sharing outside the consortium.
- ✔ The data host was mandated to sign data sharing agreements with external partners according to a standard template that is part of the CA.

Worst case scenario

- ⚑ The first company data was uploaded to the data host, but joint analyses across companies and academic data sets proved to be cumbersome because the data standards used were deviating at the detail level despite the data standards agreed on earlier in the consortium.
- ⚑ Two of the organisations also wanted to involve their U.S. colleagues in the analyses, but the GDPR data processing agreement does not allow for data sharing outside the EU.

Challenges
FAQS and strategies

CHALLENGE 5

How to share data outside of the EU in compliance with the GDPR?

CASE STUDY: A new IHI project is being proposed involving the integration of full clinical trial data (clinical + control) of previous company clinical trials.

TOPIC WRITING	PROPOSAL PREPARATION		GRANT AGREEMENT PHASE	PROJECT EXECUTION				
YEAR - 2	STAGE I — YEAR -1	STAGE II — YEAR -0,5	YEAR 0	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5

Best case scenario

✔ A sustainability strategy and plan is in place to ensure continuation after project end.

Worst case scenario

🚩 The first overarching data analyses could be completed after having improved the data harmonization across all data sets. However, there was no time left for the follow-up work as originally planned in the project.

- The companies worked out 1:1 solutions with the data providers (controllers) and their legal departments for being able to share data with the U.S. company sites. Scientists at these sites could finally also participate in the analyses.

Challenges
FAQS and strategies

CHALLENGE 3

How to best expediate data sharing within PPP?

SCENARIOS

Resources

Challenge 1: Public-private-partnerships (PPP) & Data sharing culture

IMPACT	DATA SHARING	DATA FLOW	DATA MANAGEMENT PLAN	DATA ACCESS
ARTICLE 1.1 The case of data sharing in precompetitive settings	PUBLICATION 1.3 Sharing and reuse of individual participant data from clinical trials: principles and recommendations	EXAMPLE 1.8 Schematic data flows	GUIDANCE 1.12 Data Management in EU Collaborative projects (not yet available for HE)	GUIDANCE 1.16 Data Access Committee
PUBLICATION 1.2 Status, use and impact of sharing individual participant data from clinical trials: a scoping review	PUBLICATION 1.4 Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Syntesis	EXAMPLE 1.9 Dataflow overview EXAMPLE 1.10 IMI HARMONY Data Flow	TEMPLATE 1.13 DMP template Horizon Europe	
	PUBLICATION 1.5 Sharing Is Caring – Data Sharing Initiatives in Healthcare	EXAMPLE 1.11 IMI H2O Data Flows (available from DMP)	EXAMPLE 1.14 DMP IMI H2O Project	
	EXAMPLE 1.6 Data sharing policy: example of EOSC-Life Data Sharing Policy of the COVID-19 repository		EXAMPLE 1.15 DMP IMI Conception Project	
	EXAMPLE 1.7 BigData@Heart - Responsible data sharing in a big data-driven translational research platform: lessons learned			

RESOURCES

Resources

Challenge 2: IP/Legal

INFORMED CONSENT	DATA SHARING AGREEMENTS	DATA ACCESS REQUEST	JOINT CONTROLLERSHIP AGREEMENT	IP
TEMPLATE 2.1 IMI DO-IT Project Informed Consent Forms templated	ANONYMISED EXAMPLE 2.3 IMI2 Data Transfer Agreement, including User Confirmation Sheet	ANONYMISED EXAMPLE 2.5 Anonymised example of Data Access Request Form	TEMPLATE 2.6 Joint Controllership Agreement	GUIDANCE 2.7 How does IP work?
TEMPLATE 2.2 German Medicine Informatics Initiative - broad consent	EXAMPLES 2.4 IMI EPND Material and Data Transfer Agreement			

Resources

Challenge 3: Internal Processes

INTERNAL TOOL FOR DATA RELEASE	POLICIES / PROCEDURES TO ACCESS CLINICAL TRIAL DATA	INTERNAL STRATEGIES
ARTICLE 3.1 Testimonial on internal company tool to facilitate release of data by Sean Turner	EXAMPLE 3.2 Data Sharing Policies of Vivli members	ARTICLE 3.4 Frontloading data sharing decisions in IMI/IHI projects: advantages and limitations by Eva Molero
	EXAMPLE 3.3 MSD Portal and procedure to Access to Clinical Trial Data / Public website and document Yoda	TOOL 3.5 Involvement of Roles in Data Sharing Decisions

RESOURCES

Resources

Challenge 4: Security & IT

GUIDANCE AND PRINCIPLES	DATA & IT STANDARDS	PRE-CLINICAL STANDARDS	DATA STANDARDS	OTHER
GUIDANCE 4.1 IMI Criteria for sharing data	STANDARD 4.4 ISO Information Security standards	STANDARD 4.8 SEND format for harmonized structured preclinical data	STANDARD 4.9 Fast Healthcare Interoperability Resources (FHIR)	COURSE 4.12 EHDEN 101: What is a federated data network? What is the OMOP common data model?
GUIDANCE 4.2 FAIR data principles	STANDARD 4.5 ISO/TC 215 Health informatics		GUIDANCE 4.10 Health Level 7	GUIDANCE 4.13 External guidance on the implementation of the European Medicines Agency policy on the publication of clinical data for medicinal products for human use
GUIDANCE 4.3 FAIR cookbook	STANDARD 4.6 CDISC formatting for structured data		GUIDANCE 4.11 OMOP Common data model	GUIDANCE 4.14 Health IT Standards
	GUIDANCE 4.7 EOSC-Life Report on data standards			

RESOURCES

Resources

Challenge 5: IP/GDPR

ROLES & RESPONSIBILITIES	ANONYMISATION PSEUDONYMISATION	GDPR DATA REUSE	DATA TRANSFER TO U.S.	DATA PROTECTION IMPACT ASSESSMENT
GUIDANCE 5.1 European Data Protection Board Guidelines on the concepts of controller and processor in the GDPR - dives deep on the definitions and consequences associated to the roles of Controller, Processor and Joint Controller	GUIDANCE 5.3 Introduction to anonymisation	GUIDANCE 5.7 TRANSCELERATE - A Privacy Framework for Clinical Data Reuse: Secondary Data Use in the Pharmaceutical Industry	ARTICLE 5.8 The CJEU judgment in the Schrems II case by European Parliament	ARTICLE 5.10 Data Protection Impact Assessment
	EXAMPLE 5.4 IMI HARMONY Anonymisation Concept			
ARTICLE 5.2 Key roles in GDPR by J-W Boiten	ARTICLE 5.5 Sharing Anonymised and Functionally Effective (SAFE) Data Standard for Safely Sharing Rich Clinical Trial Data		ARTICLE 5.9 SCHREMS II Summary	
	EXAMPLE 5.6 Report on Deploying Pseudonymisation Techniques by ENISA (European Union Agency for Cybersecurity)			

RESOURCES

1.1 The case of data sharing in precompetitive settings

By Magda Chlebus, Executive Director Science Policy & Regulatory Affairs at EFPIA

Our biggest societal challenges cannot be tackled in isolation. Cooperating with competitors to multiply knowledge, skills and resources without losing competitive edge is one of such strategies.

- Cooperation in data sharing contributes to benefits on multiple levels:

1. By sharing data, the organisation contributes to increasing possibilities to new discoveries and innovative ways to combine and utilize different types of data in R&D, and to reduce the time needed for drug research and development lifecycle;
2. Organizations get access to shared pool of data that might be resource-heavy, costly (or impossible) to collect and create on their own;
3. Sharing and re-use of data (with appropriate consent) helps decrease the burden for the patient community; data is not siloed with any one company.
4. By collaborating we can harness the collective intelligence of all to find workable and sustainable standards and solutions.

- This is reflected in our sector's commitments to sharing or making data accessible, beyond pandemic management.

- There is a cost of not cooperating: delaying breakthrough discoveries, overlooking new markets, missing out on cost savings by duplicating efforts.

efpia
European Federation of Pharmaceutical Industries and Associations

- Making data available to collaborative precompetitive initiatives has direct impact on all participants, and their research and innovation activities. It unlocks or validates new science, derisks new fields of science/regulatory science and enables new R&D opportunities:

  - Within clinical research, data sharing can enhance reproducibility and the generation of new knowledge, but it also has an ethical and economic dimension;
  - Scientifically, sharing makes it possible to compare and combine the data from different studies, and to more easily aggregate it for meta-analysis as well as contributing to accelerated time to and improved diagnosis (i.e., better defined phenotypes and the natural history of a condition/disease). It also increases the power of AI/Machine learning, as any one organisations data on its own does not give enough power to fuel the algorithms;
  - Identifying novel endotypes for regulatory acceptance which allow an individualized and rational treatment; this and the validation of candidates works best when results are based on accessible data from different resources;
  - It enables hypotheses to be tested more widely and conclusions to be re-examined and verified or, occasionally, corrected.

RESOURCES

- Sharing can, therefore, increase data validity, but it also draws more value from the original research investment, as well as helping to avoid unnecessary repetition of studies.
- Ethically, data sharing provides a better way to honour the generosity of clinical trial participants and to build trust, because it increases the utility of the data they provide and thus potentially also lowering the burden for other patients (i.e., of repetitive sample and data collections).
- Data access and data sharing initiatives also help demonstrate the feasibility of new access pathways (including value based models).
- Increased visibility around collaborative efforts in data sharing can also create the foundation for lawmakers and a demand for regulatory stakeholders to advance frameworks and practices to keep up with the fast-moving environment and provide enabling (and ethically sound) guidelines for fair governance that protects the patients' privacy and needs while allowing stakeholder to leverage the shared data for innovation and breakthrough discoveries.
- Data sharing allows access to broader range of data sets for more accurate conclusions to be made. "Numbers matter" to make meaningful data analsysis and to identify findings more precisely.
- It's also good for efficiency as it helps probing or jump starting potential new areas of interests by using already available data.

• There is a growing trend and expectation for transparency and reciprocity. Participation in the data sharing initiatives places all players on the same footing as far as generation of insights is concerned. It contributes to symmetry of access and knowledge for all, public and private, stakeholders.

• Doing this in a safe sandbox environment, such as institutional public private partnerships, ensures that interests of all parties are heard.

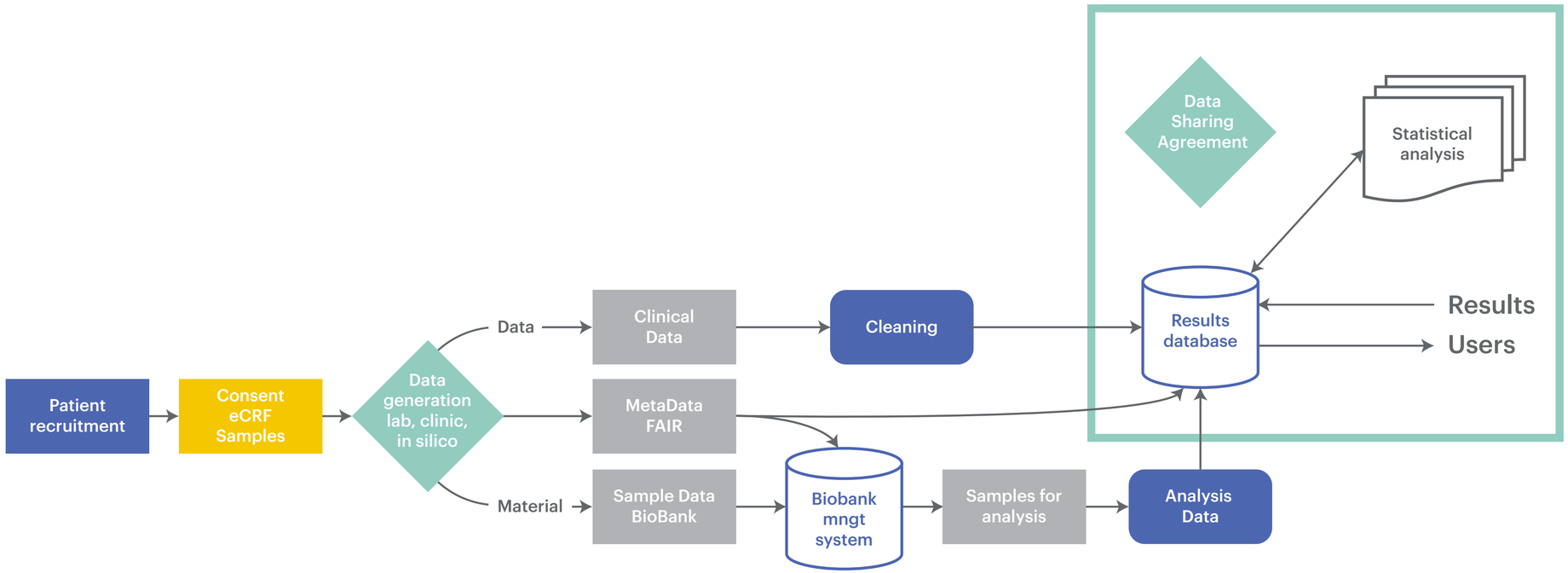
• The IMI success stories based on data sharing initiatives/platforms:

- Impact on: pediatric medicine
- Federated learning model allows competitors to pool data & identify promising molecules
- New EPND platform set to accelerate research into neurodegenerative disease
- Study of 8 million people supports neurological safety of COVID-19 vaccines
- From competition to collaboration: How secure data sharing can enable innovation

• See also:

- 1.2 Status, use and impact of sharing individual participant data from clinical trials: a scoping review
- 1.3 Sharing and reuse of individual participant data from clinical trials: principles and recommendations
- Using a Global Network of Adaptative Clinical Trials to fight COVID-19

1.8.1 Schematic data flows



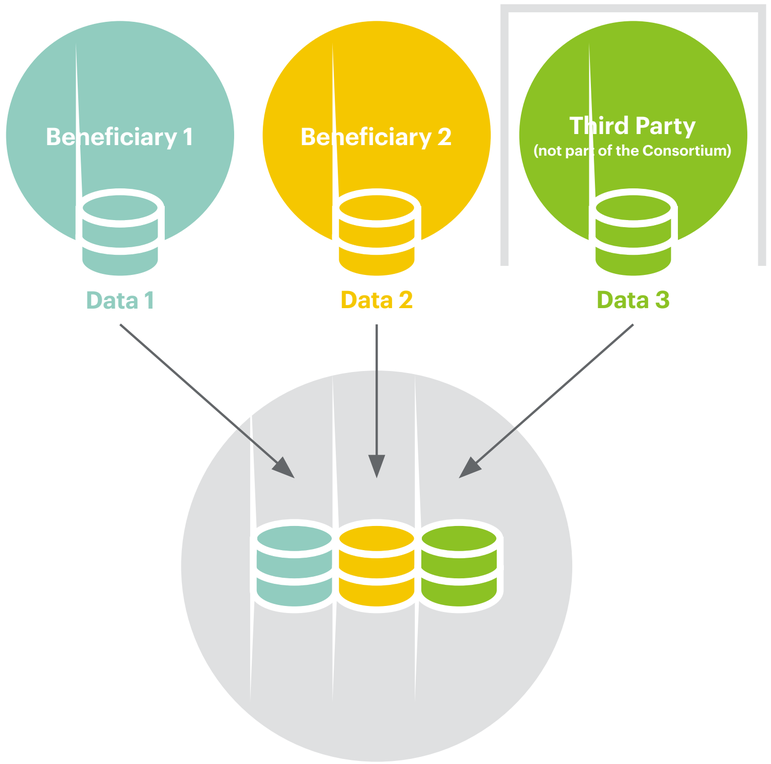
RESOURCES

1.8.2 Ways of Data Sharing in IMI/IHI Projects: Central, Federated & Hybrid

CENTRAL

Data is transferred to a central platform managed by the consortium.

The original data may be subject to further anonymisation/ pseudonymisation/harmonisation procedures [derived data].



RESOURCES

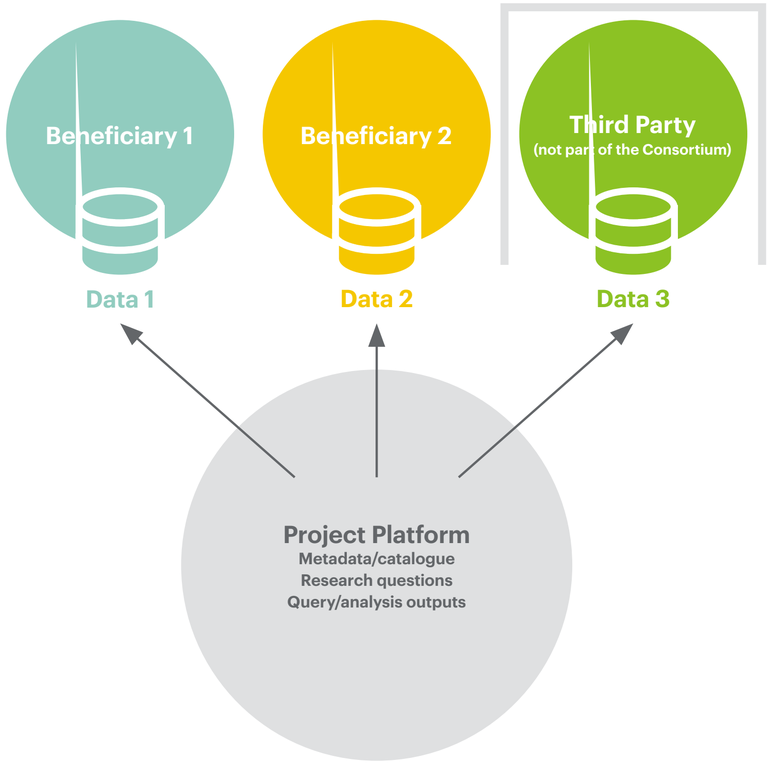
1.8.2 Ways of Data Sharing in IMI/IHI Projects: Central, Federated & Hybrid

FEDERATED

Data remains with the beneficiary/ third party contributing data and is available for querying by the consortium based on research questions posed.

Any research question is defined centrally but is still under the control of the beneficiary sharing their data.

Please note this is a general representation of a federated system.

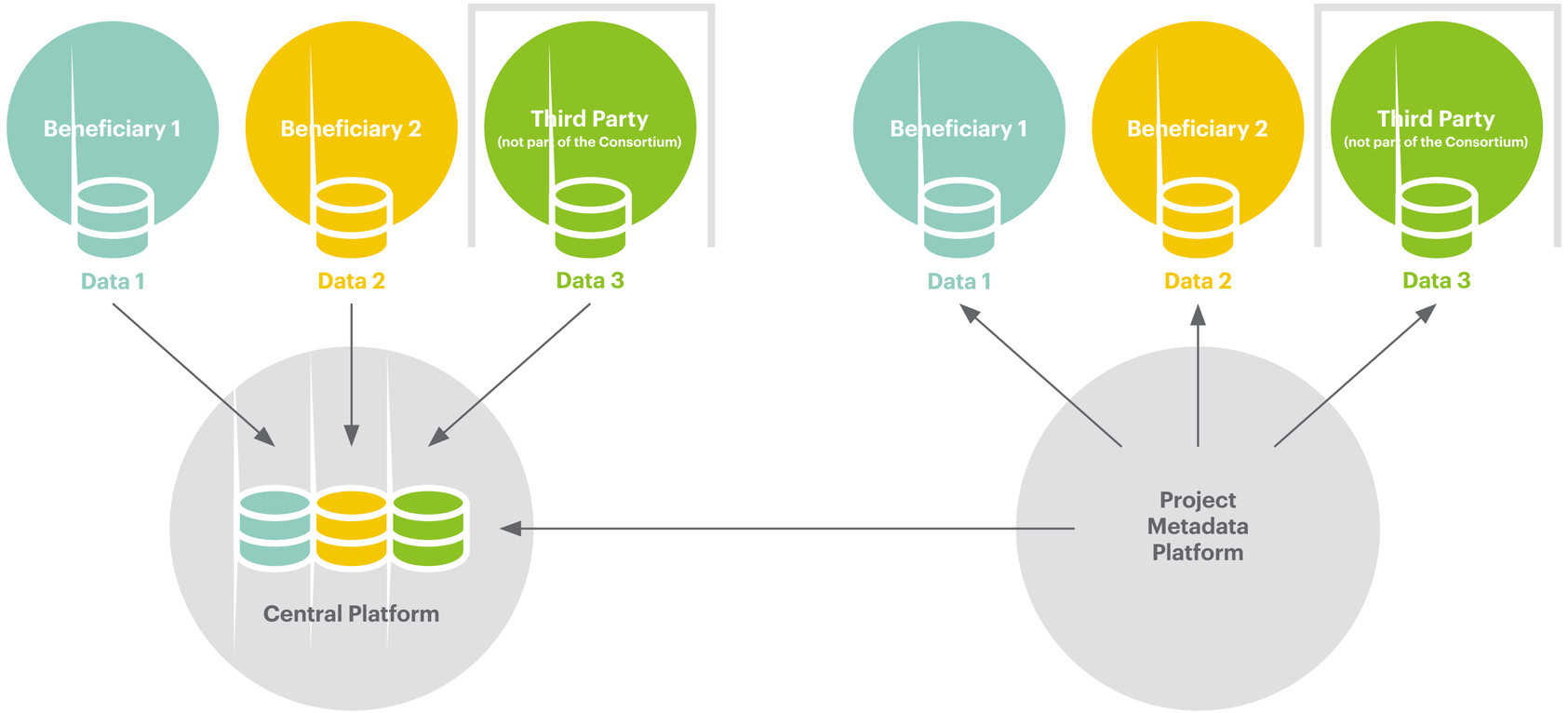


RESOURCES

1.8.2 Ways of Data Sharing in IMI/IHI Projects: Central, Federated & Hybrid

HYBRID

This model provides both central sharing and federated options for data sharing. This is one possible representation of a hybrid set-up.



RESOURCES

1.9 Data sharing options

1 Data Sharing Agreement for whole consortium Beneficiaries

All sign-up to Joint Controllership, covers all routes and can easily include data in and take data out.

- Single Data Sharing Agreement including Joint Controllership provisions.

Can be cumbersome in terms of negotiation. Might take a long time.

2 Data Sharing Agreement for bilateral arrangements between Beneficiaries

Data Transfer Agreement [DTA] in for data coming in to Host beneficiary, and DTA out with the Host. Template is provided, can share template to 'stakeholder' group for approval/comments, which may help the other partners in accepting without/with minor comments.

- A DTA is sometimes called a DSA. A DSA is also known as a Data User Access Agreement [DUAA]. Note that there can be various iterations of the name as defined by each project.

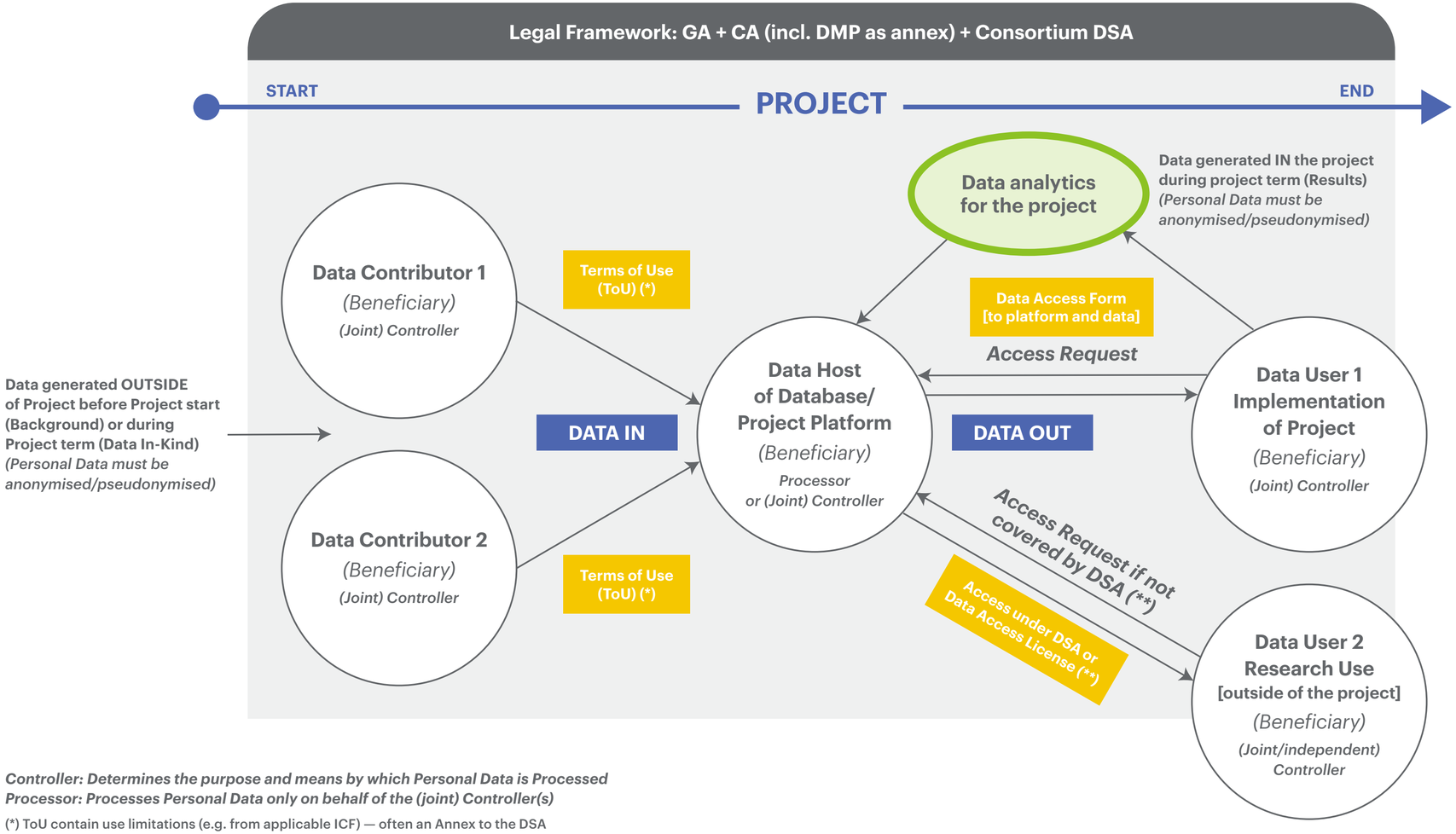
Host has to negotiate each bilateral agreement and needs to manage/align on differences in the various agreements.

If no personal identifiable information involved, a Data Transfer Record shall suffice.

RESOURCES

1.9.1 Dataflow overview

Data Sharing between Beneficiaries during Project Term



Controller: Determines the purpose and means by which Personal Data is Processed
Processor: Processes Personal Data only on behalf of the (joint) Controller(s)

(*) ToU contain use limitations (e.g. from applicable ICF) — often an Annex to the DSA
(**) Access Request for Research Use **without Research Question** → Access under consortium DSA → Data User = independent Controller
(**) Access Request for Research Use **with Research question** → Data Access license or side agreement to DSA (for agreed Purpose only) → Data User = Joint Controller
*Host might also be considered as a processor when level of decision making on pass through is determined by Data Access Policy agreed upon by all the beneficiaries

RESOURCES

3.1 Adapting an existing internal process to ease data sharing approval

By Sean Turner, Associate Director External Collaborations, AbbVie

Challenges

"At Abbvie we were faced with two added internal challenges. Firstly, the company's approach to data sharing has a strong legal/IP component. Secondly most of our central processes, that also govern our involvement in IMI projects, are headquartered in Chicago (USA), where open data sharing collaborations are less common than in Europe. Bilateral agreements with strict confidentially clauses are the norm in the USA. Both particularities complicated the release of data in IMI projects within the timeframes required. Typically, it could take over a year to get clearance through a process which was not efficient. Reviewers involved -legal expert, therapeutic lead, etc- approached each request in a siloed manner and often got lost in long chains of e-mails".

Solution

"We needed to rethink our internal procedures and ensure the translatability of processes between the USA and Europe. After looking at some potential solutions, we soon ruled out building a whole new system for data sharing. Instead, we favoured leveraging on an internal procedure that was already in place and rather familiar to us all. An existing process for the approval and release of publications, public talks, posters, etc fitted the bill very nicely.

We worked with colleagues in adapting this process, and the tools used, to include data requests needed for IMI projects. We then did several tests runs and it proved quite successful. Essentially, it is an intuitive system, accessible to everyone in the company, that centralises information and guides the flow for approval. This process typically involves around seven approvers: therapeutic area head, the requester, IP legal, compliance, etc. These are the roles needed to ensure that data -and the background behind it- can be released".

Impact

"Since implementing this new system, we have experienced a considerable improvement in data sharing efficiencies. Timescales for releasing data have been reduced considerably, from over a year to examples as short as 3 to 4 weeks for pre-clinical data sets and to 2 to 3 months for clinical data sets, depending on their sensitivity. This is indeed a very efficient internal process we have now in place which facilitates tasks, information sharing and decision making every step of the way".

RESOURCES

3.4 Frontloading data sharing decisions in IMI/IHI projects: advantages and limitations

By Eva Molero, CEO of Teamit Research

Data sharing is of critical importance in most IMI/IHI projects, but unfortunately, this is usually done via inefficient processes hampered by a diversity of issues. Altogether it often results in major project delays and repetition of the same discussions across IMI projects, leading to frustration and —more importantly— undermining the research impact potential. In our own experience and in many discussions with IMI project partners (from industry and public consortia alike), we have heard that so many issues could be avoided by simply bringing forward and frontloading the discussions and decisions that affect data sharing. If this had been done before or at project start, there would be complete clarity about the data to be shared, about the parties having responsibilities with regard to data sharing, about the data flow and the required agreements when sharing and accessing the data.

At the first glance, frontloading would be the ultimate solution to prevent risks associated to data sharing. But as usual in complex research endeavours, there is no silver bullet. It would be oversimplistic to state that thinking about those issues earlier would prevent them from happening.

First, not all data sets are fully known at the time of the project preparation and, in many cases, early project activities will determine or confirm their usefulness and suitability for the project. Secondly, in the very early stages of the project life cycle (i.e., the topic writing stage, the first stage proposal) it may be that only one part of the consortium is active (the industry consortium and the public consortium respectively) and therefore not all key stakeholders are involved in the discussions. In the second stage proposal (or single-stage IHI projects, typically the whole consortium meets for the first time, but this proposal is often developed under great time pressure leaving insufficient time for specific discussion about the minimum data sharing requirements that should be included at this stage. A similar situation is encountered during the Grant Agreement preparation. Moreover, before the start of the project all efforts invested are non-funded or non-accountable as in-kind, which disincentivises devoting significant resources to pre-project phases. And ultimately, it would not be ideal either if all data sharing aspects would be pre-decided and therefore this topic would be a closed discussion during the project implementation when new information may give rise to adaptations to the data sharing strategy.

For each project a fine balance should be sought between defining the data-sharing framework as early as possible and preparing the foundation for further refinement and detailing during the project development. For instance, having a fully-fledged data management plan before the project start may prove unrealistic, but defining the envisaged project data flow may be achievable and highly relevant. The Description of Action could then define early milestones and deliverables to report on the data sharing commitments which will facilitate dealing with those issues as early as possible in the project implementation phase.

In conclusion, frontloading high-level data sharing decisions is needed, but also designing a workplan that enables timely decisions to be taken during the project execution. Our advice: create the essential framework in the pre-project phase and define all the milestones and deliverables to grow it as the project unfolds in a staged and planned manner.

3.5 Involvement of Roles in Data Sharing decisions

- *Several Roles intervene in the process to facilitate data sharing. This swim lanes tool visually identifies who should participate in each data sharing decision or action along the project life cycle.*
- *The main decisions and actions associated with the 5 data sharing challenge areas are described in each swim lane (1- PPP and Data Sharing Processes, 2- Legal/IP, 3- Internal Processes, 4- Security/IT, 5- GDPR).*
- *In the header, the main Roles involved are represented (definitions can be found in the **Roles** section). The user can click in any of the Roles displayed and the actions in which the Role should participate will be highlighted. By ticking the "Clear selection" button the tool will be cleaned.*

CLEAR SELECTION

	IMI liaison officer	GDPR expert	IT specialist	Therapeutic lead	Senior manager/ Academic Lead	Data Protection Officer (DPO)	Lawyer	Principal investigator	Statistician	Project leader	Project coordinator
	○	○	○	○	○	○	○	○	○	○	○

	TOPIC WRITING	PROPOSAL PREPARATION	GRANT AGREMENT PHASE	PROJECT EXECUTION
CH1 PPP and Data Sharing Processes	• Outline convincing business case for data sharing • Make an initial selection of potential datasets • Ally with experienced data sharers in the consortium • Train the project leadership team	• Involve staff with previous in-depth IMI experience • Involve the consortium in selection of data sets • Consider outsourcing support activities • Draw initial data flows	• Get guidance from the IMI liaison	• Prepare internal stakeholders organise data flow workshop
CH2 Legal/IP	• Identify type of data needed to address research question • Explore legal/IP limitations of data to be shared (which should be in line with the Access Rghts in the CA)	• Examine if trial data can be shared, rights on 3rd party are adequate and informed consent is sufficient • Involve legal experts	• Map project legal components • Append Template DSA/MTA in CA • Include Mandate to sign DSA	• Confirm understanding of Data Sharing principles in consortium • Set up legal agreements asap • Develop ICF for prospective data collection
CH3 Internal Processes	• Consult internal data sharing organisational policy & processes • Perform risk/reward analysis • Identify datasets internally • Ally with experienced colleagues • Involve internal stakeholders	• Secure senior management approval • Use impact demonstrators from other projects		• Involve all internal stakeholders; be aware of staff turnover • Start data sharing clearance to expedite process • Use internal data release tools
CH4 Security/IT	• Consider security of data needed to address research question • Select auditable standard	• Utlise Common Data Model and/or Platform for Data Sharing • Set IT security requirements • Agree on data anonymisation strategy • Identify solutions to be used for Data Sharing	• Add IT security requirements to CA (e.g. via the Data Management Plan or to be included in a DSA)	• Align IT security review process across EFPIA partners • Agree to data exchange formats (to be included in the Data Management Plan, the DSA/DTA)
CH5 GDPR	• Review Initial use and access	• Check ICF template • Plan if DS outside EAA is needed • Define and agree the roles of controller, joint controller and processor	• Include DPA and/or joint controller agreement template in CA	• Develop DMP in early phase • Conduct DPIA • Organize data flow workshop checking all assumptions made • Introduce SCC for outside EAA sharing

RESOURCES

5.2 Key Roles in GDPR

By Jan Willem Boiten, Senior Program Manager at Lygature

The General Data Protection Regulation (EU) 2016/679 (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information from individuals (**"Data subjects"**) who live in the European Union (EU). It establishes the regulations on data protection and privacy, and it also addresses the transfer of personal data outside the EU. The GDPR's primary aim is to enhance individuals' control and rights over their personal data and to simplify the regulatory environment for international business.

The GDPR makes a very clear distinction between roles in the data workflow and puts different obligations on them:

- **Controller** means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. The GDPR also indicates "where two or more controllers jointly determine the purposes and means of processing, they shall be **joint controllers**". The controller could therefore be considered the "boss" in the data processing.

- **Processor** means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller. The processor could therefore be considered the instruction follower in the data processing. It is possible that the processor subcontracts another party (a "**sub-processor**"), but this is only permitted after written consent from the controller.

Obviously, the determination of these responsibilities is only critical for **personal data**; if data are not personal at all, or are fully **anonymised** then GDPR does not apply.

GDPR not only requires to clearly identify the roles of all parties involved, but also enforces that the roles and responsibilities as agreed between these parties are clearly established in legal agreements, which can take various forms depending on the actual roles in the dataflow:

1. A **data processing agreement** between the controller(s) and the processor

2. A **joint controller agreement** between controllers

3. A **controller-to-controller agreement** when the controllership of the data is transferred and the new controller of the data acts independently of the original controller, (although such an agreement is not a formal requirement defined in the GDPR text). In the context of IMI/IHI projects such a contract often takes the form of a **Data Transfer Agreement.**

Determination of the GDPR role in a contract is not legally binding; it should correspond to the facts of the processing in the project. A processor should also behave entirely according to the role set out for a processor and remain within the mandate as set forth by the controller.

It is therefore crucial to make a full data workflow analysis and determine the responsibilities in that workflow before concluding any of these agreements. Usually, these agreements also deal with liabilities in case any emergency or disaster occurs; these discussions tend to be the most difficult and time consuming. It is therefore recommended to use a standard agreement template whenever possible.

RESOURCES

Epilogue

Data sharing activities in life-science research will only grow in the next few years galvanised by current and new models of multiparty collaboration. The important roadblocks to data provision that organisations participating in IMI projects have so far experienced, will continue to expand under IHI projects unless they are ready to grasp the nettle and adopt a comprehensive and standardised approach as described in this document.

In this Data Sharing Playbook, we have focused on challenges and consensual solutions. At the same time, we underline the need for early planning and multistakeholder involvement to bypass common obstacles. By proactively considering data sharing aspects from the development of a research idea to project implementation, most bottlenecks and challenges can be prevented or significantly mitigated. In this sense, we clearly advocate for more structured, well-thought processes that can be adopted by all actors, thus unlocking the value of the data more swiftly. Organisations are encouraged to proactively adapt internal processes, develop more standardised resources and create new roles to facilitate data provision. It is to be expected that the introduction of the European Health Data Space in 2025 will underpin this approach.

Finally, the proactive, systematised approach proposed in this Playbook will only serve its purpose as long as organisations understand the full potential of data sharing and are prepared to embrace a true data sharing culture. It is our hope that this document will contribute toward this change of mindset that will significantly drive closer collaboration in ground-breaking health research.

The Authors, June 2022

This Playbook has been developed by:

Teamit Research
Eva Molero, Gisela Pairó, Martina Spadetto, Raimon Cardelús, Anna Mundet.

EATRIS
Gary Saunders, Anton Ussi, Jake Fairnie.

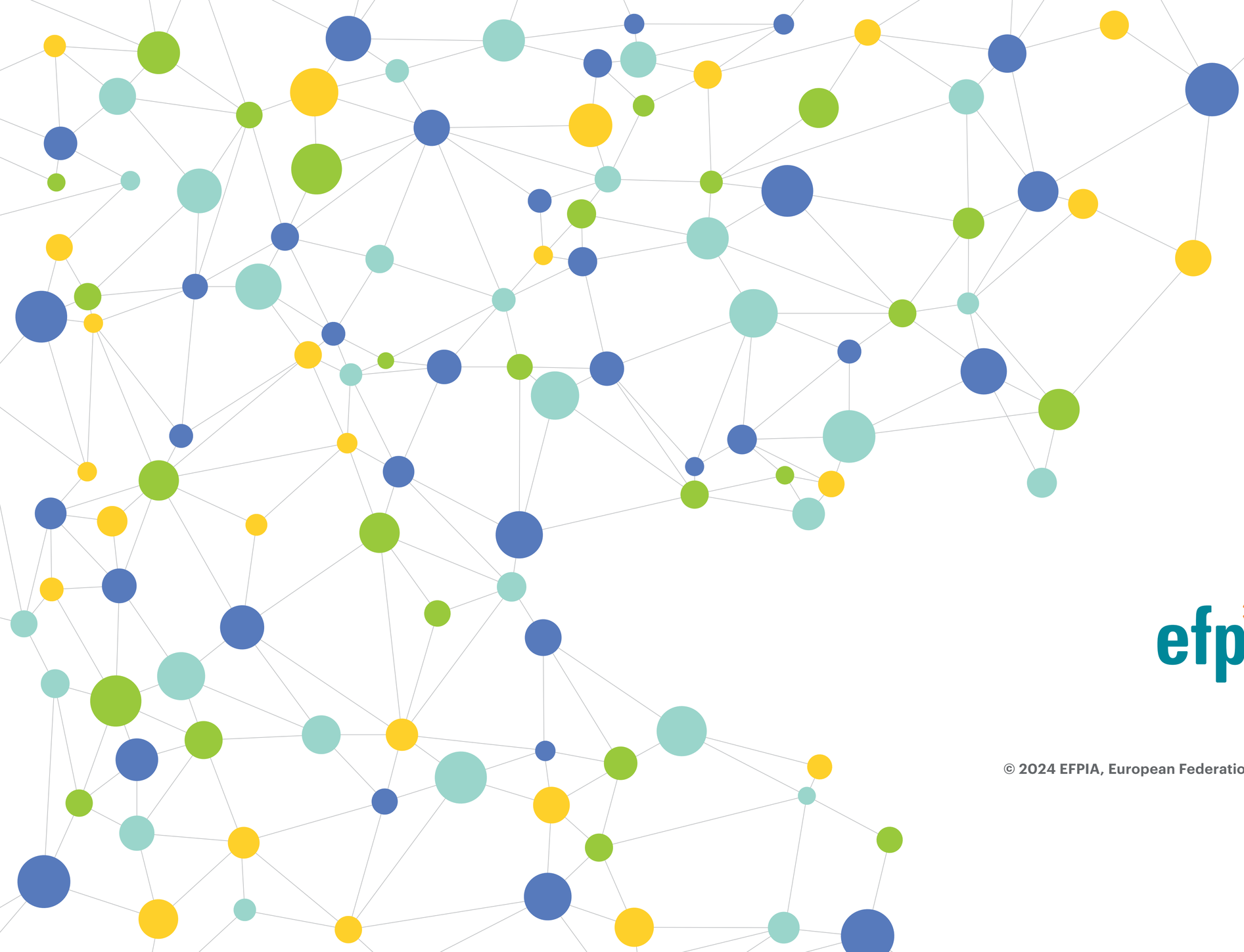
Lygature
Jan-Willem Boiten.

ITTM
Andreas Kremer.

ACKNOWLEDGEMENTS

We thank the more than 50 representatives of the EFPIA community who have contributed to the initial development of this Playbook. Their experiences and insights shared in workshops, interviews and iterative reviews have guided and enriched this Playbook.

Special thanks to Magda Chlebus, Ann Van Dessel, Karen Godbold and Brendan Barnes for their dedication and support throughout the whole process.



teamit.

eatris

lygature

ITTM
Information Technology for Translational Medicine

efpia
European Federation of Pharmaceutical Industries and Associations